PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The gender wage gap among medical specialists: A quantitative analysis of the hourly pay of publicly employed senior doctors in New Zealand
AUTHORS	Sin, Isabelle; Bruce-Brand, Bronwyn; Chambers, Charlotte

VERSION 1 – REVIEW

REVIEWER	Delattre, Eric Université de Cergy-Pontoise
REVIEW RETURNED	07-Jan-2021

GENERAL COMMENTS	Good description of the earning scheme for physicians in N.Z. Taking into account that no individual fixed effects analysis can be performed (panel data analysis), this paper makes use of the most accurate statistical technique. OLS with various controls. There is a good discussion of potential explanations for the average gap between sex. In a addition, we ought to see a- regressions 2 5 and 6 for specific specialties (of course with no specialties fixed effects) to check for the robustness among specialties (kind of Between specialties heterogeneity analysis different from table 4) b- Table 3 introduces age as a control for experience : this could lead to measurement errors which should be different between men and women. Could an IV strategy solve that potential bias ?

REVIEWER	Platt, Jonathan Columbia University, Epidemiology
REVIEW RETURNED	28-Jan-2021

GENERAL COMMENTS	This is a very nice paper, creating and leveraging a valuable data source to better understand drivers of the wage gap. The analysis is clear and straight-forward and the manuscript overall is very readable. I list a few suggested revisions below that I believe will strengthen the paper. Intro: including a final paragraph with a statement describing the paper's aims would be a helpful transition and orientation for the reader. Methods: because there were several steps in merging across data sources, it would be helpful to include a flow chart for the reader to follow who was matched and who was excluded (including the number of participants at each step) from the final analytical sample

	How were missing (non-excluded) data (e.g., marital status) modeled? Modeling 'missing' with its own dummy variable requires unrealistic assumptions and should be avoided. The authors should give more detail about missingness of the data in the analytical sample, and either multiply imputing missing data or exclude those with missing data (depending on the assumed missingness mechanisms). Instead of or in addition to the current table 1, presenting the distribution of all regression covariates by gender would allow the reader to better understand the characteristics of the sample population and compare covariate distribution between men and women. Related, my biggest concern is the gender distribution in the regression models. This is likely mitigated by the restricted nature of the sample, but I imagine that there are still significant differences in many of the model covariates. Assuming this, I would worry about how much of these estimates are based on off-support data? To some extent, this would be addressed by my previous comment, but in general, I would like the authors to address this possible limitation. Discussion: I would like to see a bit more discussion integrating this study with prior wage gap research. The wage gap in this population, while still robust, is much smaller than in the general population. What lessons can be taken from this analysis and applied more broadly to reduce the wage gap in other fields?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Eric Delattre, Université de Cergy-Pontoise
Comments to the Author:
Good description of the earning scheme for physicians in N.Z.
Taking into account that no individual fixed effects analysis can be performed (panel data analysis), this paper makes use of the most accurate statistical technique. OLS with various controls.
There is a good discussion of potential explanations for the average gap between sex.
In a addition, we ought to see
a- regressions 2 5 and 6 for specific specialties (of course with no specialties fixed effects) to check for the robustness among specialties (kind of Between specialties heterogeneity analysis different from table 4)

Great idea. We added an appendix table that presents the results of columns (2), (5), and (6) of Table 2 separately for four types of specialty: medical specialties, surgical specialties, general practice, and other specialties. The gender wage gap is fairly comparable for each of these four groups, though obviously the sample size is lower for individual groups than for the full sample, and thus the estimates of the gender wage gap are less precise.

Note we did consider running these regressions for the most disaggregated level of specialty (eg cardiologist), but the sample sizes are mostly too small to be able to draw any conclusions from these.

We added the following text describing the appendix table:

“Appendix Table 1 replicates columns (2), (5), and (6) of Table 2 separately for medical specialties, surgical specialties, general practice, and other specialties. It shows the gender wage gap is present and of comparable size for each of these specialties.”

b- Table 3 introduces age as a control for experience : this could lead to measurement errors which should be different between men and women. Could an IV strategy solve that potential bias ?

We first introduce flexible age controls as a proxy for experience in Table 2. We acknowledge age is an imperfect proxy for experience, and the relationship between age and experience might differ by gender. This is our motivation for in Table 3 attempting to better control for experience by using years since gaining medical degree and then further adjusting for estimated time away from work due to child rearing.

Although some level of measurement error in experience is likely to remain, we can’t think of an appropriate instrument that could help improve on the Table 3 estimates.

Furthermore, we note that Figure 1 shows above about age 50 there is little relationship between age (and thus experience) and earnings. This is also the age range in which the gender wage gap is largest. This lack of relationship at older ages adds further reassurance that any gender-specific mismeasurement in experience is not driving our estimated gender wage gap.

Reviewer: 2
Dr. Jonathan Platt, Columbia University
Comments to the Author:
This is a very nice paper, creating and leveraging a valuable data source to better understand drivers of the wage gap. The analysis is clear and straight-forward and the manuscript overall is very readable. I list a few suggested revisions below that I believe will strengthen the paper.

Intro: including a final paragraph with a statement describing the paper’s aims would be a helpful transition and orientation for the reader.

The last major paragraph of the introduction now reads:

Little is currently known as to the extent and drivers of gender wage gaps among senior doctors specifically. Furthermore, to the best of the authors knowledge, there are no studies examining doctor remuneration in relation to collective employment agreements or public health sector employment. The aims of this research are to quantify the gender wage gap for medical specialists in New Zealand public health system employment using actual earnings data, with a focus on controlling for factors such as experience, hours worked, and medical speciality, which are commonly ascribed factors for gender wage gaps.

Methods: because there were several steps in merging across data sources, it would be helpful to include a flow chart for the reader to follow who was matched and who was excluded (including the number of participants at each step) from the final analytical sample

This is a great idea. We added a Figure (now Figure 1) that shows how the data sets are brought together and the number of observations lost at each stage of creating the final data set.

How were missing (non-excluded) data (e.g., marital status) modeled? Modeling ‘missing’ with its own

dummy variable requires unrealistic assumptions and should be avoided. The authors should give more detail about missingness of the data in the analytical sample, and either multiply imputing missing data or exclude those with missing data (depending on the assumed missingness mechanisms).

This is a good point.

The number of observations with missing covariates in our data is very low. Although the reasons for the missing values are unclear, failing to fully complete the Census form is likely a major driver, and this may not be entirely random.

Missing covariates become an issue only from column 6 of Table 2. In our revision, we now drop the 108 observations with missing covariates from this specification and Tables 3 and 4. To illustrate that the assumptions required for this do not affect our results, we alternatively impute all missing values to maximise the gender wage gap and then impute them to minimise the gender wage gap. We find the coefficient on female in column (6) of Table 2, which is -0.119 in the specification that drops missing observations, varies only from -0.117 to -0.121. We conclude our treatment of missing values has minimal effect on our findings. For subsequent tables we present results dropping missings only.

Instead of or in addition to the current table 1, presenting the distribution of all regression covariates by gender would allow the reader to better understand the characteristics of the sample population and compare covariate distribution between men and women.

Great idea. We now present summary statistics in Table 1 separately by gender.

Related, my biggest concern is the gender distribution in the regression models. This is likely mitigated by the restricted nature of the sample, but I imagine that there are still significant differences in many of the model covariates. Assuming this, I would worry about how much of these estimates are based on off-support data? To some extent, this would be addressed by my previous comment, but in general, I would like the authors to address this possible limitation.

As part of addressing your previous point, we adjusted the variables presented in Table 1 to more closely align with the exact controls used in our regressions. It is now straightforward to see how much overlap there is in the covariate ranges of male and female specialists. Most of our covariates are indicator variables, and in nearly every case both male and female doctors exhibit a good distribution across categories. The variable that differs most between genders may be hours worked in other jobs, with women less likely to work high hours. However, men are only three times as likely as women to be in the highest hours category.

Given the large overlap in these distributions, we don’t believe lack of common support is a concern.

Discussion:
I would like to see a bit more discussion integrating this study with prior wage gap research. The wage gap in this population, while still robust, is much smaller than in the general population. What lessons can be taken from this analysis and applied more broadly to reduce the wage gap in other fields?

We now discuss in the discussion section how the wage gap we estimate compares to the wage gap in New Zealand on average for high-skill occupations, and suggest reasons why it is lower. In particular, we suggest the MECA has a role in limiting negotiation, and that the shortage of doctors limits the ability of DHBs to gender discriminate in their hiring and pay.

VERSION 2 – REVIEW

REVIEWER	Platt, Jonathan Columbia University, Epidemiology
REVIEW RETURNED	15-Mar-2021

GENERAL COMMENTS	The authors have fully addressed my comments.